

Analisis Multidimensional

En un mercado tan competitivo como el que enfrentan las organizaciones actuales, la información se ha convertido en el activo más importante. No solo como secreto comercial, sino como indicador clave para tomar las decisiones adecuadas en los momentos oportunos. La información es el principal activo de una empresa.

Cada vez que un analista pretende realizar el estudio del comportamiento del negocio, es inherente encontrar más de una perspectiva que permita contextualizar los datos, es decir existe mas de una variable que nos indique desde una perspectiva dada la situación de la organización. Por ejemplo podemos analizar las ventas de la empresa por Producto, o por territorio, o por Cliente y Almacén, etc. Podríamos igualmente analizar las ventas, los costos, los stocks y las deudas para cada una de las variables anteriores.



Figura 1

Esta situación es completamente normal, ya que son muchas las variables que afectan el entorno. Así se encuentran variables como el cliente, país, producto, tiempo, canal de distribución, cuenta contable,

escenario, venta neta, impuesto, descuento, costo, y un infinito etcétera de variables que es imposible enumerar acá.

De esta necesidad de analizar la empresa desde distintas vistas o perspectivas y sobre diversos aspectos de rendimiento o valores, surge el concepto de **análisis multidimensional**. El cual es definido como la capacidad de contextualizar una o más variables (medidas como por ejemplo las ventas, costos, stocks, deudas, etc.) a través del empleo de perspectivas (dimensiones como los productos, clientes, filiales, almacenes). Este proceso de contextualización se lleva a través de estructuras multidimensionales, que en palabras simples son asociaciones de tablas y de valores numéricos bajo un modelo de diseño de estrella. En el modelo la tabla de valores o hechos está en el centro y las tablas que contextualizan los valores o que ofrecen diversas perspectivas de los valores están alrededor de esta tabla central, y se denominan dimensiones. Las medidas generalmente serán numéricas y las dimensiones generalmente serán alfanuméricas. Sin embargo, pueden presentarse excepciones a esta regla.



Figura 2

El análisis multidimensional basado en procesos OLAP (On Line Application Program) tiene como eje central las estructuras multidimensionales, también denominadas cubos o datamarts. Un cubo es una estructura multitabular (compuesta por diversas tablas) modelada a través del esquema estrella en la que se puede apreciar el centro conformado por una tabla de valores que serán agregados y en las puntas las

tablas de atributos o perspectivas de analisis. El cubo es el eje central sobre el que gira todo el analisis multidimensional.

En una estructura multidimensional o cubo las dimensiones son componentes estructurales de estas y estan conformadas de atributos que en algunos casos mantienen una relación de jerarquia. Por ejemplo la tabla de dimensiones de producto contiene los atributos:Clase, Linea, estilo, etc. La jerarquia de atributos es una estructura relacionada de tipo padre a hijo, por ejemplo en la dimensión de producto, el estilo es hijo del atributo Linea, este es es hijo del atributo clase,y asi sucesivamente.

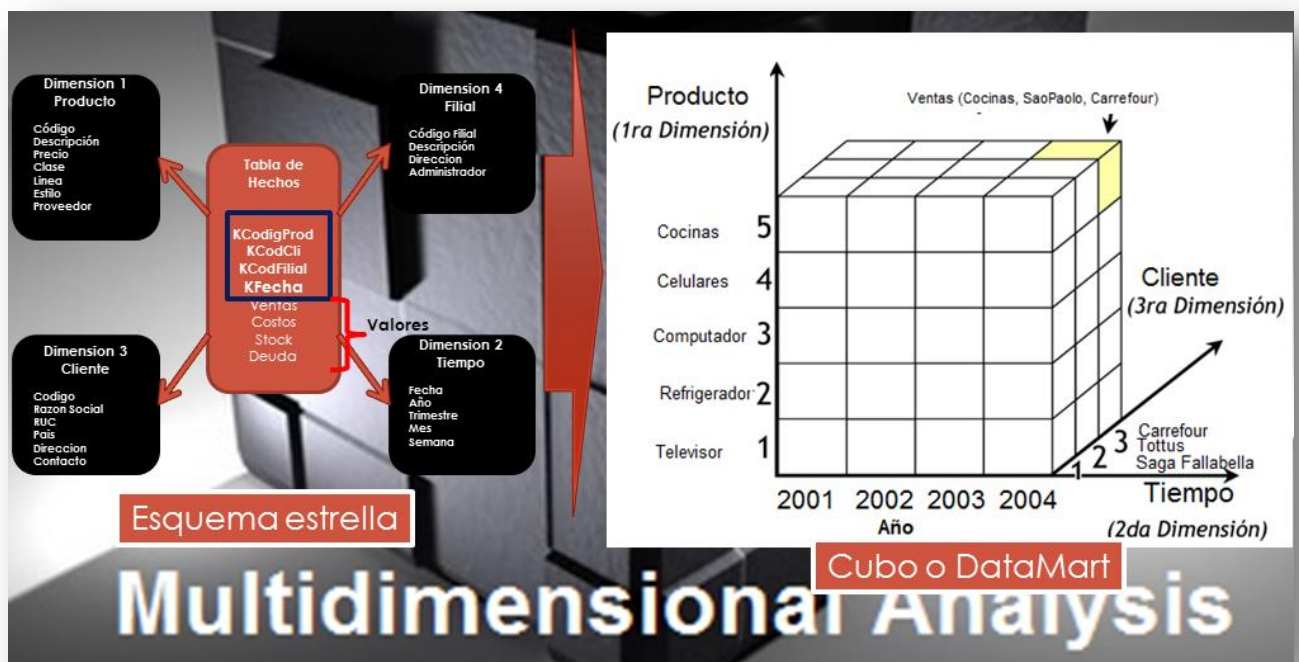


Figura 3

El analisis OLAP.es un proceso inherente a las plataformas de BI que permite a los analistas crear una gran variedad de escenarios analiticos a la velocidad del pensamiento de enorme utilidad para los decisores, a través de la visualización de modelos de variables numericas asociadas a diversas perspectivas o dimensiones utilizando la técnica de drag and drop y procedimientos de filtros.

Una vez definida la estructura multidimensional o cubo, a traves de la interface de diseño podemos modelar distintos informes analiticos. Con la tecnica de selección y arrastre (drag and drop) podemos seleccionar los valores a mostrar, y las perspectivas o dimensiones a partir de las cuales los valores numericos seran agregados,.

En este ejercicio podemos ver como analizamos las ventas y los costos de los productos por region de ventas y lineas de productos. Para esto arrastramos los valores de venta y costos al area respectiva y las regiones en el lado de las filas y la linea de productos en el lado de las columnas.

El area de filtros nos permite seleccionar los años del analisis en este caso 2004, asi como tambien el grupo territorial, en este caso Europa.

Podemos cambiar los ejes de filas y columnas para ver desde otro angulo el mismo informe. Tambien podemos incorporar las clases a nuestro analisis y mover todas estas dimensiones de acuerdo a nuestras necesidades.

Por ejemplo el mismo criterio lo podemos emplear para analisis de ventas y costos por año y trimestres de ventas, simplemente arrastrando los objetos de datos de nuestro cubo.

Una de las ventajas de las herramientas de Business Intelligence es la alta calidad grafica y visual de los modelos asi como su interaccion con el usuario. En el análisis multidimensional los resultados pueden cambiarse dinamicamente para mostrar diversas perspectivas con las tecnicas o funciones de drag and drop y filtros que proporcionan la interface. El análisis multidimensional permite modificar las perspectivas del informe analítico, a la velocidad del pensamiento mientras que las funcionalidades de filtro permiten trabajar con contenidos más específicos.

Características del Modelo Multidimensional

En general, la estructura básica de un DW para el Modelo Multidimensional está definida por dos elementos: esquemas y tablas.

- **Tablas DW:** como cualquier base de datos relacional, un DW se compone de tablas. Hay dos tipos básicos de tablas en el Modelo Multidimensional:
 - **Tablas Fact o tablas de hechos:** contienen los valores de las medidas de negocios, por ejemplo: ventas promedio en dólares, número de unidades vendidas, cantidad del stock de un producto, etc.
 - **Tablas Lock_up o tablas de dimensiones:** contienen el detalle de los valores que se encuentran asociados a la tabla Fact. Es decir son los campos a los que los valores estan referenciando. Por ejemplo: **Las ventas promedios en dolares pueden referirse a:** toda la compañía, un producto especifico, un departamento, una región o un país.



- **Esquemas DW:** la colección de tablas en el DW se conoce como Esquema. Los esquemas caen dentro de dos categorías básicas: esquemas estrellas y esquemas snowflake.

Conceptos asociados al Datawarehouse

Tabla de Hechos (Tablas Fac)

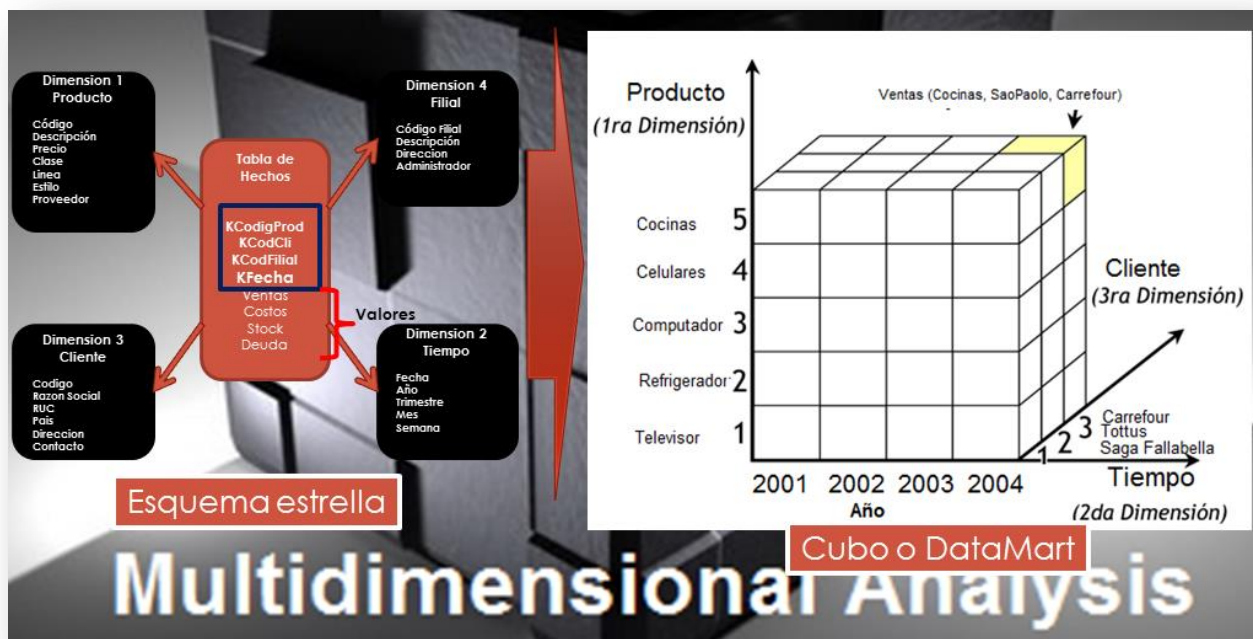
Es la tabla central en un esquema dimensional. Es en ella donde se almacenan las mediciones numéricas del negocio. Estas medidas se registran a un nivel de detalle o granularidad (unidad básica de la tabla).

La granularidad de la tabla queda determinada por el nivel de detalle que se almacenará en la tabla. Por ejemplo, para el caso de la dimension tiempo, podemos almacenar el detalle los valores numericos a nivel de año, trimestre, mes, semana, la granularidad puede ser el importe vendida anualmente, por trimestre, por mes, por semana, etc..

La clave de la tabla fact recibe el nombre de clave compuesta o concatenada debido a que se forma de la composición (o concatenación) de las llaves primarias de las tablas dimensionales a las que está unida.

Así entonces, se distinguen dos tipos de columnas en una tabla fact: columnas fact y columnas key.

Donde la columna fact es la que almacena alguna medida de negocio y una columna key forma parte de la clave compuesta de la tabla.



Tablas Dimensionales (Tablas Lock-up)

Estas tablas son las que se conectan a la tabla fact. Una tabla lock_up almacena un conjunto de atributos que están relacionados a una dimensión particular.

Tablas lock_up no contienen hechos, en su lugar los valores en las tablas lock_up son los elementos que determinan la estructura de las dimensiones.

Una tabla lock_up está compuesta de una primary key que identifica unívocamente una fila en la tabla junto con un conjunto de atributos, y dependiendo del diseño del modelo multidimensional puede existir una foreign key que determina su relación con otra tabla lock_up.

Para decidir si un campo de datos es un atributo o un hecho se analiza la variación de la medida a través del tiempo. Si varía continuamente implicaría tomarlo como un hecho, caso contrario será un atributo.

Los atributos dimensionales son un rol determinante en un DDW. Ellos son la fuente de todas las necesidades que debieran cubrirse. Esto significa que la base de datos será tan buena como lo sean los atributos dimensionales, mientras más descriptivos, manejables y de buena calidad, mejor será el DDW.